

UNITED STATES PATENT APPLICATION
FOR
CONTENT ADDRESSABLE MEMORY WITH RANGE COMPARE FUNCTION

Attorney Docket No.: N1-P104

Inventor: Jose P. Pereira

Prepared By:
Charles Shemwell, Attorney At Law
998 East El Camino Real, Suite 204
Sunnyvale, California 94087-7913
Tel.: 408-736-3221 Fax: 408-732-3248

EXPRESS MAIL CERTIFICATE OF MAILING

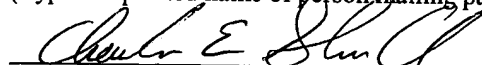
"Express Mail" mailing label number: EL 910 085 107 US

Date of Deposit: September 24, 2001

I hereby certify that this paper is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

Charles E. Shemwell

(Typed or printed name of person mailing paper(s) or fee(s))



(Signature of person mailing paper or fee)

CONTENT ADDRESSABLE MEMORY WITH RANGE COMPARE FUNCTION

FIELD OF THE INVENTION

The present invention relates generally to content addressable memory devices, and more particularly to comparison operations within content addressable memory devices.

5 BACKGROUND

Content addressable memory (CAM) devices are often used to perform packet forwarding and classification operations in network switches and routers. A CAM device can be instructed to compare a comparand value, typically formed from one or more fields within the header of an incoming packet, with CAM words stored in an associative storage array within the CAM device. If the comparand matches a CAM word, the CAM device generates a CAM index that identifies the location of the matching CAM word within the storage array, and asserts a match flag to signal the match. The CAM index may then be used to index another storage array, either within or separate from the CAM device, to retrieve routing or classification information for the packet.

Compare operations within a CAM device are typically carried out on a bit-by-bit basis to determine if the comparand is identical to a CAM word. In some cases, selected bits within the comparand or the CAM word may be masked, enabling a range of comparands to match the CAM word. Unfortunately, the range of matching comparands achieved by masking is restricted in significant ways. First, the total range of matching comparands is limited to a power of two value (i.e., the inclusive difference between the highest and lowest matching comparands is 2^N , where N is the number of contiguous masked bits). A second restriction of a mask-based range is that the lower limit of the range generally must be an integral multiple of the total range (i.e., lower limit = 0×2^N , 1×2^N , 2×2^N , ...). If a match is desired for a range of comparands that is not a power of two or for which the lower limit is not an integral multiple of the total range, then

multiple compare operations typically must be performed, each compare operation covering a portion of the overall range and meeting the above mask-based range restrictions. Referring to Fig. 1, for example, to determine whether a given comparand has an eight-bit source port field within the range 23 to 46, six different compare operations are needed (i.e., comparisons against values 23 and 46, and against mask-based ranges 31-24, 39-32, 43-40, and 45-44 as shown in Fig. 1). Similarly, six different compare operations are needed to determine whether a sixteen-bit destination port field within a comparand exceeds 1023 (i.e., comparisons against mask-based ranges 2047-1024, 4095-2048, 8191-4096, 16383-8192, 32767-16384, and 65535-32768 as shown in Fig. 1).

The multiple compare operations used to compare a comparand against a range of values may be performed successively, for example using different global mask values to define the different mask-based ranges, or they may be performed simultaneously by comparing a comparand to multiple CAM words in the CAM array, each CAM word constituting a range component that covers a portion of the overall range. Simultaneous comparison against multiple range components (i.e., single-cycle range compare) is faster than performing successive compare operations, but consumes a relatively large amount of CAM array storage. In the case of a binary CAM device, for example, each value in a given range must typically be stored in the CAM array to enable a single-cycle range compare. For instance, to permit the source port to range from 23 to 46 as shown in Fig. 1, 24 range components (i.e., 24 CAM words) would need to be stored in the CAM array. Alternatively, in the case of a ternary CAM device (i.e., a CAM device that includes storage for local mask values to permit per-row masking of CAM words), single-cycle range compare may be enabled by storing one range component in the CAM array for each mask-based range (setting the local mask value for the entry to define the mask-based range) and one range component for each of the exact values in the range. Referring to Fig. 1,

for example, six range components may be stored in a ternary CAM device to cover the source port range from 23 to 46: exact value 23; mask-based range 31-24; mask-based range 39-32; mask-based range 43-40; mask-based range 45-44, and exact value 46.

The consumption of CAM array storage due to storage of multiple range components becomes more severe if more than one field within a comparand is permitted to range.

Typically, the total number of range components that must be stored is the product of the numbers of range components required to define each distinct range. Referring to Fig. 1, for example, if a match is desired on any comparand for which the source port is between 23 to 46 (a range defined by 6 range components) and for which the destination port is greater than 1023 (a range also defined by 6 range components), then a total of 36 range components (i.e., 6x6) must be stored in a typical ternary CAM array to fully define the range of comparands which will match the source port and destination port ranges. The number of range components proliferates as other fields within the comparand are permitted to range (e.g., fields such as the source address, destination address, protocol, type of service, and virtual private network, etc.), consuming further storage space within the CAM array.

SUMMARY

A content addressable memory (CAM) device having one or more ranging CAM cells is disclosed in numerous embodiments. In one embodiment, each ranging CAM cell includes a pair of storage elements to store an upper bound value and a lower bound value, respectively, and a compare circuit to compare a comparand value against the upper bound value and the lower bound value. The compare circuit outputs a greater-than signal if the comparand is above the upper bound and a less-than signal if the comparand is below the lower bound. By chaining a set of ranging CAM cells together such that each ranging CAM cell in the chain except the least significant is coupled to receive the greater-than/less-than signals from a less significant ranging CAM cell in the chain, the individual upper bound values and lower bound values are combined to define a desired range of matching comparands. The greater-than and less-than signals output by the most significant ranging CAM cell in the chain may then be coupled to a match line to signal an out-of-range condition if an incoming comparand is determined to be greater than or less than the search range.

In alternative embodiments, instead of providing a pair of storage elements within each ranging CAM cell to define a search range having an upper and lower bound, a single storage element may be used within each ranging CAM cell to define an upper or a lower bound. In such embodiments, a compare circuit is provided to compare a comparand value against the upper or lower bound to generate an output indicative of whether the bound is exceeded or not.

In yet other embodiments, multi-mode CAM cells within the CAM array may be selectively switched between a ranging mode of operation and either a binary or ternary mode of operation.

These and other features and advantages of the present invention are described in the detailed description below.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings and in which like reference numerals refer to similar elements and in which:

5 Fig. 1 illustrates a combination of exact and mask-based values that may be stored in a prior art CAM device to cover a desired range of source port and destination port values;

 Fig. 2 illustrates a ranging CAM device according to an embodiment of the present invention;

 Fig. 3 illustrates an embodiment of a ranging CAM array and its connections to a comparand register and read/write circuit;

 Fig. 4 illustrates an exemplary row of CAM cells coupled to a match line;

 Fig. 5 depicts a chain of four ranging CAM cells that have exemplary upper bound and lower bound bits stored therein;

 Fig. 6 illustrates a chain of ranging CAM cells according to an embodiment for storing an inclusive explicit range definition;

 Fig. 7 is a logic diagram of an upper bound compare circuit;

 Fig. 8 is a logic diagram of a lower bound compare circuit;

 Fig. 9 illustrates an embodiment of a least significant ranging CAM cell that can be used in a chain of ranging CAM cells to store an exclusive range definition;

20 Fig. 10 illustrates the use of range definitions to express an exemplary classification rule within a single entry in a ranging CAM array;

 Fig. 11 illustrates a multi-mode CAM cell according to one embodiment;

 Fig. 12 illustrates a mode configuration circuit that may be used to set the operating modes of multi-mode CAM cells within a ranging CAM array;

Fig. 13 illustrates an embodiment of a ranging CAM array that includes multiple independently addressable and searchable blocks of CAM cells;

Fig. 14 illustrates a multi-mode CAM cell that may be switched between an implicit range mode and a binary mode;

5 Fig. 15 illustrates an alternative compare circuit that may be used in place of the binary compare circuit of Fig. 14;

Fig. 16 illustrates a multi-mode CAM cell according to another alternative embodiment;

Fig. 17 illustrates a system that includes a host processor and a ranging CAM device; and

Fig. 18 illustrates the operation of the host processor of Fig. 17 according to one embodiment.

10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100

DETAILED DESCRIPTION

Ranging CAM Device

Fig. 2 illustrates a ranging CAM device 100 according to an embodiment of the present invention. The ranging CAM device 100 includes a ranging CAM array 101, address circuit 106, instruction decoder 108, flag generator 107, priority encoder 105, comparand register 104 and read/write circuit 103. Instructions, addresses and commands are input to the CAM device 100 via an instruction bus 111, address bus 113 and comparand bus 115, respectively. Each of the buses is preferably a multi-conductor signal path coupled to at least one host device, such as a general purpose processor, digital signal processor, network processor, application specific integrated circuit (ASIC) or other instruction issuing device. Also, in alternative embodiments, one or more of the buses may be eliminated and the corresponding signals time-multiplexed onto another of the buses.

The ranging CAM array 101 is coupled to (i.e., connected directly to or through one or more intervening circuits) the address circuit 106, priority encoder 105, flag generator 107, comparand register 104, and read/write circuit 103. The address circuit 106 is used to select a particular row of the ranging CAM array for read or write access. The read/write circuit 103 is used to sense the output of the selected row during a read operation and to transmit a value to the selected row during a write operation. The comparand register 104 is used to store a comparand value received via the comparand bus 115, and outputs the comparand value to the CAM array 101. In alternative embodiments the comparand register 104 may be omitted and the comparand value input directly to the ranging CAM array 101 from the comparand bus 115.

The ranging CAM array 101 includes a plurality of CAM cells, some or all of which may be ranging CAM cells according to embodiments described below. Each row of CAM cells is used to store one or more data words referred to herein as CAM words. Each row of CAM cells

may additionally include storage for a validity value that indicates whether a valid CAM word is stored within the row. In the case of segmented rows, discussed below, a separate validity value may be stored for each row segment.

During a compare operation, the comparand may be masked by a global mask value, then compared simultaneously with all the CAM words stored in the ranging CAM array 101. Each of the rows of CAM cells in the ranging CAM array 101 is coupled to a respective match line 127, and any match between the comparand and a valid CAM word results in a match signal being output to the priority encoder 105 and the flag generator 107 via the match line 127. When one or more match signals are asserted on the match lines 127, the priority encoder 105 selects one of the match signals and outputs a CAM index 174 (i.e., address of the CAM word corresponding to the selected match signal). The flag generator 107 also receives the match signals, and outputs a match flag signal (MF) to indicate that a match has occurred. If more than one match signal is asserted, the flag circuit may additionally output a multiple match flag signal (MMF) to indicate that multiple matches have occurred.

Each CAM word stored within the ranging CAM array 101 may include one or more exact values, bit-masked values, and/or range definitions. An exact value is a pattern of bits that represents a single numeric value. A bit-masked value is an exact value in which one or more bits are masked (e.g., by a local mask value in a ternary CAM cell), thereby specifying a power-of-two range. In the case of a compare operation, exact values and bit-masked values are similar in that both types of values are compared for equality with an incoming comparand. That is, a match between a comparand and an exact or bit-masked value occurs when the comparand is equal to the exact or bit-masked value bit-for-bit after any bit masking.

Range definitions are values that define bounded numeric ranges. In contrast to exact values and bit-masked values, the value (or values) that form a range definition are not compared

for equality with an incoming comparand, but rather are used to define the upper and lower bounds of an inequality expression. A comparand that falls within the bounded numeric range is deemed to match the range definition. Range definitions may be either explicit or implicit. An explicit range definition is a vector quantity that includes two component values, an upper bound value and a lower bound value. Together, the upper bound value and the lower bound value define a range of comparands that will be deemed to match the range definition. An explicit range definition may be inclusive or exclusive of the upper and lower bound values. In the case of an inclusive explicit range definition, a match is signaled if the following expression is true:

$$\text{lower bound} \leq \text{comparand} \leq \text{upper bound}.$$

In the case of exclusive explicit range definition, a match is signaled if the following expression is true:

$$\text{lower bound} < \text{comparand} < \text{upper bound}.$$

An implicit range definition includes only one bound value, upper or lower, with the other bound impliedly being the maximum or minimum value that can be numerically represented by the comparand value (or portion thereof). For example, in the case of a sixteen bit field within a CAM word (and corresponding comparand), an implicit range definition may include a lower bound value of 1024 to define a range between 1024 and 65535 (65535 being the maximum possible value for the sixteen-bit field). Any comparand value of 1024 or greater (i.e., up to 65535) falls within the bounded range defined by the threshold value and is said to "match" the implicit range definition. Like explicit range definitions, implicit range definitions can include or exclude the bound value.

Ranging CAM Array

Fig. 3 illustrates an embodiment of the ranging CAM array 101 and its connections to the comparand register 104 and read/write circuit 103. As shown, the ranging CAM array 101

includes a plurality of CAM cells 301 arranged in rows and columns. Each of the CAM cells is designated "T/R" to indicate that it may be a ternary CAM cell, a ranging CAM cell, or a multi-mode CAM cell that may be selectively operated in either a ternary mode or a ranging mode.

Each of the CAM cells may also be a binary CAM cell as discussed below. Each row of the

5 CAM cells is coupled to a respective word line 125 and to a respective match line 127. Note that one or more additional word lines 125 may be provided to enable independent access to different sets of storage elements within CAM array 101. For example, a first word line may be activated to access (e.g., write or read) data storage elements within a row of ternary CAM cells, while a second word line may be activated at the same time or in a pipelined fashion to access mask storage elements within the same or a different row of ternary CAM cells. CAM array 101 may also include more than one match line 127 per row of CAM cells. For example, multiple compare circuits may be provided within all or a portion of the CAM cells 301 to enable pipelined or parallel compare operations, each compare circuit being coupled to a respective one of multiple match lines 127.

Each column of the CAM cells is coupled to the comparand register 104 via a pair of comparand signal lines (C and \bar{C}), and to the read/write circuit 103 via a first pair of bit lines (B and \bar{B}) and via a second pair of bit lines (MB and \overline{MB}). (One or more additional pairs of comparand lines may be provided per column of CAM cells to enable pipelined compare operations.) During a compare operation, each pair of comparand signal lines is used to provide
20 a comparand signal (e.g., a bit from the comparand value stored in the comparand register 104) and its complement to each CAM cell 301 in the corresponding column such that the comparand is provided, bit-for-bit, to each column of CAM cells in the ranging CAM array 101. Each CAM cell within each row of CAM cells generates a compare result which is used to affect the logical state of the match line 127. Although each of the CAM cells 301 is depicted in Fig. 3 as being

coupled to a corresponding match line, one or more ranging CAM cells may not be coupled to a match line 127 and may instead be coupled to output a compare result to another ranging CAM cell to form a chain of ranging CAM cells.

In one embodiment, each of the CAM cells 301 includes first and second storage elements that may be used to store, in the case of a ternary CAM cell, a data value and a local mask value, or, in the case of a ranging CAM cell, an upper bound value and a lower bound value. In the case of a multi-mode CAM cell, the first and second storage elements are used to store either data and mask or upper and lower bounds according to the mode of operation selected for the CAM cell. In alternative embodiments, each of the CAM cells may include only a single storage element as discussed below. During a read or write access to the CAM array 101, one of the word lines 125 is activated by an address circuit (e.g., element 106 of Fig. 1) to select one of the rows of CAM cells for read or write access. In a read operation, each CAM cell in the selected row is enabled to output the contents of the first storage element to the read/write circuit 103 via the first pair of bit lines (e.g., as a complimentary pair of signals) and to output the contents of the second storage element to the read/write circuit 103 via the second pair of bit lines. Conversely, during a write operation, each CAM cell in the selected row is enabled to store in the first storage element a value received from the read/write circuit 103 via the first pair of bit lines, and to store in the second storage element a value received from the read/write circuit 103 via the second pair of bit lines. Because distinct pairs of bit lines are provided to carry data to and from the first and second storage elements within each CAM cell of a selected row, the contents of both storage elements within the CAM cells of a selected row may be read or written simultaneously. Also, because dedicated pairs of comparand signal lines are provided to input a comparand value to the ranging CAM array 101, a compare operation may be performed simultaneously (or at least concurrently, that is, at least partly overlapping in time)

with reading or writing a selected row of CAM cells. In alternative embodiments, only one pair of bit lines may be provided per column of CAM cells for time-multiplexed access (read or write) to the first and second storage elements within a given row of CAM cells. Similarly, the dedicated pairs of comparand signal lines may be omitted and the bit lines time multiplexed (i.e., between read/write and compare operations) to supply comparand values to the CAM array 101. In the embodiment of Fig. 3, the CAM cells 301 within each row are depicted as being coupled to the corresponding match line 127 in a wired OR combination (e.g., to force a precharged match line to a particular logic level in the event of a mismatch). Alternatively, a NAND structure may be used to logically combine the match results serially before affecting the logic state of the match line 127. For example, pass gates having gate terminals coupled to the respective match outputs of the CAM cells 301 within a given row may be coupled in series with one another between a match line output and a reference voltage so that, if any of the CAM cells 301 outputs a mismatch signal, the corresponding pass gate will be switched to a non-conducting state to interrupt the connection between the match line and the reference voltage, thereby signaling the mismatch condition (e.g., by causing the match line output to float or be charged by a pull up or pull down connection). More generally, any logical structure may be used to generate a match (or mismatch) signal for a row of CAM cells without departing from the spirit and scope of the present invention.

Ranging CAM Cell

Fig. 4 illustrates an exemplary row of CAM cells 400 coupled to a match line 127. The row of CAM cells 400 includes a set of ternary CAM cells, 401(T-1)-401(0), and a set of ranging CAM cells, 403(R-1)-403(0), T and R representing respective integer values that may range from one to any number. The ternary CAM cells 401(T-1)-401(0) may be used to store an exact value or a bit-masked value, and the ranging CAM cells 403(R-1)-403(0) may be used to store an

explicit or implicit range definition. More or fewer sets of ternary CAM cells and more or fewer sets of ranging CAM cells may be included in the row 400 in alternative embodiments. Also, binary CAM cells may be used instead of or in addition to the ternary CAM cells 401(T-1)-401(0).

5 The ranging CAM cells 403(R-1)-403(0) are coupled to one another in a chain that is ordered from a least significant ranging CAM cell 403(0) to a most significant ranging CAM cell 403(R-1). The most significant CAM cell 403(R-1) outputs a greater-than (GT) signal, GT_{R-1} , and a less-than (LT) signal, LT_{R-1} , on respective signal lines that are coupled in a wired-OR configuration to the match line 127. When asserted, the greater than signal, GT_{R-1} , indicates that a comparand value has been determined to be greater than an upper bound value stored within the row of CAM cells 400. Similarly, the less-than signal, LT_{R-1} , indicates that a comparand value has been determined to be less than a lower bound value stored within the row of CAM cells 400. Each of the less significant ranging CAM cells 403(R-2)-403(0)) is not directly connected to the match line, but rather is coupled to the match line through transistors in one or more of the more significant CAM cells and outputs a GT/LT signal pair to the next more significant ranging CAM cell in the chain. In one embodiment (i.e., an embodiment for storing an exclusive range definition), each ranging CAM cell 403 outputs an active low GT signal if either of the following conditions is met:

(CG1) a comparand value is greater than an upper bound value stored within the ranging
20 CAM cell; or

(CG2) the comparand value is equal to the upper bound value, and a GT signal from a next less significant ranging CAM cell is asserted (i.e., set to a low logic state).

Similarly, each ranging CAM cell 403 outputs an active low LT signal if either of the following conditions is met:

(CL1) the comparand value is less than a lower bound value stored within the ranging CAM cell; or

(CL2) the comparand value is equal to the lower bound value, and a LT signal from a next less significant ranging CAM cell is asserted (i.e., set to a low logic state).

5 Note that neither the second condition for assertion of the GT signal nor the second condition for assertion of the LT signal will be met for the least significant ranging CAM cell 403(0) because there is no less significant ranging CAM cell to supply a GT or LT signal. As discussed below, the least significant ranging CAM cell 403(0) can be alternatively configured to effect storage of either an inclusive or exclusive range definition within the chain of ranging CAM cells 403(R-1)-403(0).

10 The GT and LT signals generated by a given ranging CAM cell 403 may be expressed analytically as follows:

15 $\overline{GT}_i = (C_i > U_i) \text{ OR } (C_i = U_i \text{ AND } \overline{GT}_{i-1}); \text{ and}$

20 $\overline{LT}_i = (C_i < L_i) \text{ OR } (C_i = L_i \text{ AND } \overline{LT}_{i-1})$, where 'i' is an integer value indicating the position of the ranging CAM cell in the chain (0 to R-1), U is the upper bound, L is the lower bound and C is the comparand value. Note that the '—' symbol is used in reference to GT and LT signals herein and in the accompanying figures to indicate that, in the exemplary embodiments shown, the GT and LT signals are active low signals. In alternative embodiments, the GT and LT signals and any other signals described herein may have active logic states opposite those shown.

 Assuming that the comparand value received within each ranging CAM cell is a bit, and that the upper and lower bounds are also bits, the above expressions may be expressed in Boolean form as follows:

$$\overline{GT}_i = (C_i \bullet \overline{U}_i) + (C_i \bullet \overline{GT}_{i-1}) + (\overline{U}_i \bullet \overline{GT}_{i-1}); \text{ and}$$

$$\overline{LT}_i = (\overline{C}_i \bullet L_i) + (\overline{C}_i \bullet \overline{LT}_{i-1}) + (L_i \bullet \overline{LT}_{i-1}), \text{ where the '+' symbol indicates a logical}$$

OR operation and the '•' symbol indicates a logical AND operation.

In the embodiment of Fig. 4, referred to herein as a wired-NOR embodiment, the match line is precharged to a high logical level by a pre-charge circuit 405. Any one or more of the ternary CAM cells 401(T-1)-401(0) may signal a mismatch by pulling the match line 127 to a low logic level. Similarly, the chain of ranging CAM cells may signal a mismatch (i.e., a comparand out of range condition) by outputting a logic low GT_{R-1} signal or LT_{R-1} signal to pull the match line 127 to a low logic level. In an alternative embodiment, match indications from the individual ternary CAM cells 401(T-1)-401(0) may be combined with one another and with a match indication from the chain of ranging CAM cells 403(R-1)-403(0) in a logical AND arrangement such that, if each of the ternary CAM cells 401 indicates a match, and if the chain of ranging CAM cells 403 indicates a match, a match signal will be asserted (either active high or active low) for the CAM row or segment thereof.

Fig. 5 depicts a chain of four ranging CAM cells 503(3)-503(0) that have exemplary upper bound and lower bound bits stored therein. Together, the four upper bound bits define the upper bound value 1001b ('b' indicating binary notation), or decimal 9. The four lower bound bits define the lower bound value 0101b, or decimal 5. Thus, the upper and lower bound values constitute an explicit range definition from 5 to 9 within a four-bit field. Consider the operation of the ranging CAM cells 503 when compared against a comparand in which a corresponding four-bit field (i.e., $C_{i+3}-C_i$, where i is an integer value indicating the offset of the field within the comparand) has the value 0111b (decimal 7). Applying the conditions CG1 and CG2 set forth above (and assuming that the least significant ranging CAM cell 503(0) is configured to produce an inclusive range definition), the least significant ranging CAM cell 503(0) will deassert the

GT₀ signal (e.g., by allowing GT₀ to float or by outputting a logic high GT₀ signal) because the least significant bit of the comparand field, C_i, (a logic '1') is not greater than the upper bound value, U₀ (also a '1'). The next more significant ranging CAM cell 503(1) will assert the GT₁ signal because C_{i+1}, a '1,' is greater than upper bound value U₁, a '0.' Similarly, ranging CAM cell 503(2) will assert the GT₂ signal because C_{i+2}, a '1,' is greater than upper bound value U₂, a '0.' The most significant ranging CAM cell 503(3), however, will deassert the GT₃ signal because C_{i+3}, a '0,' is not greater than the upper bound value U₃, a '1,' (thereby preventing condition CG1 from being met), and because C_{i+3} is not equal to the upper bound value U₃ (thereby preventing condition CG2 from being met).

Applying conditions CL1 and CL2 set forth above to compare the comparand field to the lower bound value, the least significant ranging CAM cell 503(0) will deassert the LT₀ signal because C_i is not less than L₀. The next more significant ranging CAM cell 503(1) will deassert the LT₁ signal because C_{i+1} is greater than L₁, thus preventing either condition CL1 or CL2 from being met. Ranging CAM cell 503(2) will deassert the LT₂ signal because C_{i+2} is not less than L₂ (preventing condition CL1 from being met) and because signal LT₁ is not asserted, thereby preventing condition CL2 from being met. Similarly, the most significant ranging CAM cell 503(3) will deassert the LT₃ signal because C_{i+3} is not less than L₃ and because signal LT₂ is not asserted. Finally, because neither the LT₃ signal nor the GT₃ signal is asserted, the match line 127 is not pulled down by the chain of ranging CAM cells 503. Accordingly, the chain of ranging CAM cells 503 signals a match, at least as to the comparand field C_{i+3}-C_i.

Still referring to Fig. 5, consider the operation of the ranging CAM cells 503 if the comparand field C_{i+3}-C_i has a value of 1010b (decimal 10). Ranging CAM cell 503(1) will assert the GT₁ signal because C_{i+1} is greater than U₁. Ranging CAM cell 503(2) will, in turn, assert the GT₂ signal because C_{i+2} is equal to U₂ and the GT₁ signal is asserted. Finally, ranging CAM cell

503(3) will assert the GT_3 signal because C_{i+3} is equal to U_3 and the GT_2 signal is asserted. The asserted GT_3 signal (i.e., logic low GT_3 signal) will pull the match line 127 down to indicate that the comparand field does not match the range definition. The mismatch indication is as expected because the comparand value 10 is outside the range 5 to 9.

5 As another example, consider the operation of the ranging CAM cells 503 if the comparand field $C_{i+3}-C_i$ has a value of 0100b (decimal 4). Ranging CAM cell 503(0) will assert the LT_0 signal because C_i is less than L_0 . Ranging CAM cell 503(1) will, in turn, assert the LT_1 signal because C_{i+1} is equal to L_1 and the LT_0 signal is asserted. Similarly, ranging CAM cell 503(2) will assert the LT_2 signal because C_{i+2} is equal to L_2 and the LT_1 signal is asserted. Finally, ranging CAM cell 503(3) will assert the LT_3 signal because C_{i+3} is equal to L_3 and the LT_2 signal is asserted. The asserted LT_3 signal (i.e., logic low LT_3 signal) will pull the match line 127 down to indicate that the comparand field does not match the range definition. Again, the mismatch indication is as expected because the comparand value 4 is outside the range 5 to 9.

Reflecting on the operation of ranging CAM cells 503, note that in a typical ternary CAM device, the following three CAM words would need to be stored in the CAM array to allow single-cycle comparison of the four-bit comparand field against the range 5 to 9:

- 1) exact value 5 (0101b);
- 2) bit-masked range 6-7 (011xb, the 'x' indicating a masked bit); and
- 3) and bit-masked range 8-9 (100xb).

20 By contrast, using the CAM cell chain of Fig. 5, only a single CAM word is needed to define the range 5 to 9. Note that a specific field size and range of numeric values have been shown in Fig. 5 for the purpose of example only. Numerous different numeric ranges may be defined within fields of more or fewer bits without departing from the spirit or scope of the present invention.

Fig. 6 illustrates a chain of ranging CAM cells 603(R-1)-603(0) according to an embodiment for storing an inclusive explicit range definition. Each ranging CAM cell 603 includes a pair of storage elements 605 and 615 to store an upper bound value and lower bound value, respectively. In one embodiment, each of the storage elements 605, 615 is a static storage element implemented as a bi-stable latch (e.g., back-to-back coupled inverters). In alternative embodiments, different types of storage elements may be used including, without limitation, dynamic storage elements (typically implemented by a single transistor and charge storage element), non-volatile storage elements or any other type of storage element that can be used to store digital data. Except for the least significant ranging CAM cell 603(0), each ranging CAM cell 603 includes an upper bound compare circuit 607 and a lower bound compare circuit 617 (each depicted as a dashed line box within CAM cell 603(R-1)). The upper bound compare circuit 607 includes a first pair of transistors 602 and 604 disposed to form a logic NAND gate having an output coupled to the GT output of the ranging CAM cell 603 and having first and second inputs coupled respectively to receive a comparand signal from a comparand line and to receive an upper bound value from the upper bound storage element 605. More specifically, a drain terminal of transistor 602 is coupled to the GT output of the CAM cell 603, a source terminal of transistor 602 is coupled to the drain terminal of transistor 604, and a gate terminal of transistor 602 is coupled to the comparand line to receive a comparand value C. A gate terminal of transistor 604 is coupled to the upper bound storage element to receive an active low version of the upper bound value (i.e., \overline{U}), and a source terminal of transistor 604 is coupled to a reference voltage (e.g., ground). Note that additional circuitry may be interposed between the source terminal of the transistor 604 and the reference voltage to provide for timing control of a range compare operation. For example, an additional transistor, controlled by a timing signal, may be coupled between the transistor 604 and a ground node such that, when the timing signal

is asserted, the source terminal of the transistor 604 is coupled to the ground reference potential, thereby enabling the NAND function of the transistor pair 602, 604.

The upper bound compare circuit 607 further includes a second pair of transistors 606 and 608 disposed to form a logic OR gate having an output coupled to the GT output of the ranging CAM cell 603 and having first and second inputs coupled respectively to receive the comparand signal and to receive the upper bound value. More specifically, a drain terminal of transistor 606 is coupled to the GT output of the ranging CAM cell 603 and to a drain terminal of transistor 608. A source terminal of transistor 606 is coupled to a GT input of the ranging CAM cell 603 (i.e., to receive the GT signal output by a less significant ranging CAM cell) and to a source terminal of transistor 608. A gate terminal of transistor 606 is coupled to the gate terminal of transistor 604 and therefore to the upper bound storage element 605. A gate terminal of the transistor 608 is coupled to the gate terminal of transistor 602 and therefore to the comparand line.

The upper bound compare circuit 607 implements the following Boolean expression:

$$\overline{GT}_i = (C_i \bullet \overline{U}_i) + ((C_i + \overline{U}_i) \bullet \overline{GT}_{i-1}), \quad (1)$$

which, upon expansion of the second term, becomes the sum of products discussed above, namely:

$$\overline{GT}_i = (C_i \bullet \overline{U}_i) + (C_i \bullet \overline{GT}_{i-1}) + (\overline{U}_i \bullet \overline{GT}_{i-1}). \quad (2)$$

Fig. 7 is a logic diagram of the upper bound compare circuit 607 and corresponds to the Boolean expression (1), above. More specifically, the comparand value C_i is ANDed with a complemented upper bound value \overline{U}_i in logic AND gate 701. Logic AND gate 701 outputs a signal 702 which, if set, indicates that C_i is greater than U_i ; the condition CG1 set forth above for asserting the active low \overline{GT}_i signal. The comparand value C_i is also ORed with the

complemented upper bound value \overline{U}_i in logic OR gate 703. The output of the logic OR gate 703 is set if C_i is greater than or equal to \overline{U}_i . Consequently, if signal 702 is low (indicating that C_i is not greater than U_i), then a high output from logic OR gate 703 indicates that C_i is equal to U_i .

The output of logic OR gate 703 is ANDed with the active low GT signal from the next less

5 significant CAM cell (designated \overline{GT}_{i-1} in Fig. 7) in logic AND gate 705 to produce signal 704.

Signal 704, if high while signal 702 is low, indicates that C_i is equal to U_i and that the GT signal from the next less significant ranging CAM cell is asserted; the condition CG2 above for

asserting the GT_i signal. The signals 702 and 704 are ORed in logic NOR gate 707 to produce

the active low GT_i signal (designated \overline{GT}_i in Fig. 7). Accordingly, if either signal 702 or 704 is asserted, indicating that either condition CG1 or condition CG2 is met, the GT_i signal is asserted.

Note that, in the wired NOR embodiments shown in Figs. 4-6, the GT signal is active low.

Alternative logic states may be used in other embodiments without departing from the spirit and scope of the invention.

Returning to Fig. 6, each ranging CAM cell (except the least significant) also includes a lower bound compare circuit 617. The lower bound compare circuit 617 includes transistors 612, 614, 616 and 618 coupled in the same manner as transistors 602, 604, 606 and 608, respectively, of the upper bound circuit 607 except that the gate terminals of transistors 612 and 618 are coupled to receive a complemented version of the comparand signal \overline{C} , the gate terminals of transistors 614 and 616 are coupled to the lower bound storage element to receive a lower bound value L, the drain terminals of the transistors 612, 616 and 618 are coupled to the LT output of the CAM cell instead of the GT output, and the source terminals of the transistors 616 and 618 are coupled to an LT input of the CAM cell (instead of the GT input) to receive an

LT signal from a less significant ranging CAM cell. The lower bound compare circuit 617 implements the following Boolean expression:

$$\overline{LT}_i = (\overline{C}_i \bullet L_i) + ((\overline{C}_i + L_i) \bullet \overline{LT}_{i-1}), \quad (3)$$

which, upon expansion of the second term, becomes the sum of products discussed above,
5 namely:

$$\overline{LT}_i = (\overline{C}_i \bullet L_i) + (\overline{C}_i \bullet \overline{LT}_{i-1}) + (L_i \bullet \overline{LT}_{i-1}). \quad (4)$$

Fig. 8 is a logic diagram of the lower bound compare circuit 617 and corresponds to the Boolean expression (3), above. More specifically, the complemented comparand value \overline{C}_i is ANDed with the lower bound value L_i in logic AND gate 801. Logic AND gate 801 outputs a signal 802 which, if set, indicates that C_i is less than L_i ; the condition CL1 set forth above for asserting the LT_i signal. The complemented comparand value \overline{C}_i is also ORed with the lower bound value L_i in logic OR gate 803. The output of the logic OR gate 803 is set if C_i is less than or equal to L_i . Consequently, if signal 802 is low (indicating that C_i is not less than L_i), then a high output from logic OR gate 803 indicates that C_i is equal to L_i . The output of logic OR gate 803 is ANDed with the active low LT signal from the next less significant CAM cell (designated \overline{LT}_{i-1} in Fig. 8) in logic AND gate 805 to produce signal 804. Signal 804, if high while signal 802 is low, indicates that C_i is equal to L_i and that the LT signal from the next less significant ranging CAM cell is asserted; the condition CL2 above for asserting the LT_i signal. The signals 802 and 804 are ORed in logic NOR gate 807 to produce the active low LT_i signal (designated \overline{LT}_i in Fig. 8). Accordingly, if either signal 802 or 804 is asserted, indicating that either condition CL1 or condition CL2 is met, the LT_i signal is asserted. Note that, in the wired NOR embodiments shown in Figs. 4-6, the LT signal is active low. Alternative logic states may be used in other embodiments without departing from the spirit and scope of the invention.

Returning to the embodiment of Fig. 6, the least significant ranging CAM cell 603(0) differs from the other ranging CAM cells 603 in that no GT input signal is received.

Accordingly, an upper bound compare circuit 609 within the ranging CAM cell 603(0) includes transistors 602 and 604 to compare the comparand value C_0 to the upper bound value U_0 , but omits transistors 606 and 608 (which are used in the more significant ranging CAM cells to pass the state of the input GT signal to the GT output if the comparand value is equal to the upper bound value). As in the more significant ranging CAM cells, transistors 602 and 604 implement a logic NAND circuit to drive (i.e., pull) the GT_0 output low if the comparand value C_0 is greater than the upper bound value U_0 .

The ranging CAM cell 603(0) also does not receive an LT input signal. Accordingly, a lower bound compare circuit 619 within the ranging CAM cell 603(0) includes transistors 612 and 614 to compare the comparand value C_0 to the lower bound value L_0 , but omits transistors 616 and 618 (which are used in the more significant ranging CAM cells to pass the state of the input LT signal to the LT output if the comparand value is equal to the lower bound value). As in the more significant ranging CAM cells, transistors 612 and 614 implement a logic NAND circuit to drive (i.e., pull) the LT_0 output low if the comparand value C_0 is less than the lower bound value L_0 .

Reflecting on the operation of the chain of CAM cells 603(R-1)-603(0), it can be seen that if the input comparand value (or selected field of bits therein) is identical to the overall upper bound value or lower bound value stored within the chain, the GT_{R-1} and LT_{R-1} signals will be deasserted (e.g., allowed to float or driven high), thereby signaling a match condition. The resulting inclusive range definition (i.e., range that includes bounding values) occurs because the GT_0 and LT_0 signals, which pass through each of the more significant ranging CAM cells to drive the GT_{R-1} and LT_{R-1} signals, respectively, are deasserted.

Fig. 9 illustrates an embodiment of a least significant ranging CAM cell 903 that can be used in place of ranging CAM cell 603(0) of Fig. 6 to allow storage of an exclusive range definition. The CAM cell 903 includes upper and lower bound storage elements 905 and 915 coupled respectively to upper and lower bound compare circuits 909 and 919. The upper bound compare circuit 909 includes transistors 902 and 904, each having a drain terminal coupled to the GT output of the CAM cell 903 and each having a source terminal coupled to a reference voltage (e.g., coupled to ground directly or through a timing control element). A gate terminal of transistor 902 is coupled to receive an inverted version of the upper bound value (i.e., \overline{U}_0) from the upper bound storage element 905, and a gate terminal of the transistor 904 is coupled to receive a comparand value C_0 from a comparand signal line. By this arrangement, the GT_0 signal is asserted (pulled low in this example) if the comparand value C_0 is high, or if the upper bound value U_0 is low. That is, the GT_0 signal is asserted if the comparand value C_0 is greater than or equal to the upper bound value U_0 .

The lower bound compare circuit 919 includes transistors 912 and 914, each having a drain terminal coupled to the LT output of the CAM cell 903 and each having a source terminal coupled to a reference voltage (e.g., coupled to ground directly or through a timing control element). A gate terminal of the transistor 912 is coupled to receive the lower bound value L_0 from the lower bound storage element 915, and a gate terminal of the transistor 914 is coupled to receive a complemented comparand value (i.e., \overline{C}_0) from a comparand signal line. By this arrangement, the LT_0 signal is asserted (pulled low in this example) if the comparand value C_0 is low, or if the lower bound value L_0 is high. That is, the LT_0 signal is asserted if the comparand value C_0 is less than or equal to the lower bound value L_0 .

Referring to Figs. 6 and 9, consider the effect of substituting CAM cell 903 for CAM cell 603(0) within the chain of ranging CAM cells 603. If an input comparand, or field thereof, is

bit-for-bit identical to an upper bound value stored in the chain of ranging CAM cells, CAM cell 903 will assert the GT_0 signal. Each of the more significant CAM cells in the chain 603 will forward the asserted GT_0 signal to the next more significant CAM cell in the chain such that the GT_0 signal will ultimately drive the GT_{R-1} signal to a logic low state, signaling a mismatch.

- 5 Similarly, if the comparand is bit-for-bit identical to a lower bound stored in the chain of ranging CAM cells, the CAM cell 903 will assert signal LT_0 , and each of the more significant CAM cells in the chain 603 will forward the asserted LT_0 signal to the LT_{R-1} output, again signaling a mismatch. Thus, if CAM cell 903 is included in the least significant position of a ranging CAM cell chain, a mismatch will be signaled by the chain if the comparand value is equal to either the upper or lower bound values stored within the chain. Accordingly, in such an embodiment, the upper and lower bound values stored within the chain of ranging CAM cells constitute an exclusive explicit range definition.

Although ranging CAM cells have been described in reference to Figs. 4-9 as outputting both a GT and a LT signal, in alternative embodiments, a chain of CAM cells may generate only a GT signal or only a LT signal. Each CAM cell in such a chain may be implemented as described in reference to Fig. 6, but in a first case omitting the lower bound storage element and lower bound compare circuit, and in a second case omitting the upper bound storage element and upper bound compare circuit. In the first case, the chain of CAM cells may be used to store an implicit range definition having a stored upper bound and an implied lower bound (i.e., $00...0b$, the lowest possible value for the corresponding field of bits within a comparand value). In the second case, the chain of CAM cells may be used to store an implicit range definition having a stored lower bound and an implied upper bound (i.e., $11...1b$, the maximum possible value for the corresponding field of bits within a comparand value). In an alternative embodiment, the least significant CAM cell in the chain may be replaced by the CAM cell 903 of Fig. 9 (omitting

either the lower bound storage element 915 and compare circuit 919, or the upper bound storage element 905 and compare circuit 909) so that an exclusive implicit range definition may be stored in the chain of CAM cells instead of an inclusive implicit range definition.

Fig. 10 illustrates the use of range definitions to express an exemplary classification rule within a single entry in a ranging CAM array. Under the classification rule (which may be specified, for example, by a human operator), an incoming Internet Protocol (IP) packet having (1) a source port value between 192.11.07.23 to 192.11.07.46, inclusive, and (2) a destination port value greater than 1023 is to be classified as a class 0 packet.

To implement the source port component of the classification rule, the exact value 192.11.07 (each '.' delimited field within the source port value being an eight-bit value) is stored within a set of 24 ternary CAM cells 1001 within a row of CAM cells 1000, and an inclusive explicit range definition having an upper bound value of 46 and a lower bound value of 23 is stored in a chain of eight ranging CAM cells 1003 within the row of CAM cells 1000. If a comparand ranging from 192.11.07.23 to 192.11.07.46 is compared against the CAM word, the set of ternary CAM cells will indicate a match as to the exact value 192.11.07, and the chain of ranging CAM cells 1003 will indicate a match for the range from 23 to 46. To implement the destination port component of the classification rule (the destination port field being a sixteen-bit field), a chain of sixteen ranging CAM cells 1005 is used to store an inclusive explicit range definition having an upper bound value of 65535 and a lower bound value of 1024.

Consequently, any comparand having a destination port value between 1024 and 65535, inclusive, will match the range definition. In an alternative embodiment, only the lower bound value of 1024 need be stored in the chain of ranging CAM cells 1005, with the upper bound value of 65535 being implied (i.e., an implicit range definition may be stored). Accordingly, the

upper bound storage element and upper bound compare circuitry may be omitted from the chain of ranging CAM cells 1005.

As demonstrated by the classification example of Fig. 10, providing multiple distinct chains of ranging CAM cells within the row of CAM cells 1000 enables multiple range compare operations to be performed simultaneously on respective fields of a comparand, yet requires storage of only a single CAM word. As discussed above in Fig. 1, a typical ternary CAM array requires storage of thirty-six CAM words to produce the same single-cycle comparison result. Additional chains of ranging CAM cells may be used to store range definitions for other fields within incoming comparand values, potentially saving even more CAM array storage.

Multi-Mode CAM Cell

As mentioned above, multi-mode CAM cells are CAM cells that may be selectively operated in different modes. In one embodiment, for example, a multi-mode CAM cell may be switched between a ternary mode of operation and a ranging mode of operation. In another embodiment, a multi-mode CAM cell may be switched between a binary mode of operation and a ranging mode of operation. In yet another embodiment, a multi-mode CAM cell may be switched between different ranging CAM cell modes.

Fig. 11 illustrates a multi-mode CAM cell 1103 according to one embodiment. The multi-mode CAM cell 1103 includes a pair of storage elements 1150 and 1160, a range compare circuit 1105, a ternary compare circuit 1109, and a mode select input to receive a mode select signal 1102. The state of the mode select signal 1102, high or low, selects either a ternary or ranging mode of operation for the multi-mode CAM cell 1103 by enabling either the range compare circuit 1105 or the ternary compare circuit 1109 and disabling the other.

The significance of the data stored in the storage elements 1150 and 1160 depends upon the operating mode of the multi-mode CAM cell 1103. When the CAM cell 1103 is in the

ternary mode, the storage element 1150 is used to store a data value and the storage element 1160 is used to store a mask value. By contrast, when the CAM cell 1103 is in the range mode, the storage element 1150 is used to store an upper bound value and the storage element 1160 is used to store a lower bound value. In the embodiment of Fig. 11, the storage elements 1150 and 1160 are implemented by back-to-back inverters, but, as discussed in reference to Fig. 6, numerous other types of storage cells may be used in alternative embodiments.

The range compare circuit 1105 includes an upper bound compare circuit to generate a GT signal (GT_i) and a lower bound compare circuit to generate a LT signal (LT_i). The upper bound compare circuit is implemented by transistors 602, 604, 606 and 608, which are coupled to one another and coupled to receive a complemented upper bound value \bar{U} from storage element 1150, an incoming GT signal (GT_{i-1}), and a comparand value C, all as described above in reference to Fig. 6. The lower bound compare circuit is implemented by transistors 612, 614, 616 and 618, which are coupled to one another and coupled to receive a lower bound value L from storage element 1160, an incoming LT signal (LT_{i-1}), and a complemented comparand value \bar{C} , also as described in reference to Fig. 6. Unlike the embodiment of Fig. 6, however, the source terminals of transistors 604 and 614 are coupled to the drain terminal of a range mode enable transistor 1107, rather than directly to a reference voltage. A source terminal of the range mode enable transistor 1107 is coupled to a ground reference voltage (additional circuitry such as timing control circuitry may be interposed between the source terminal of transistor 1107 and the reference voltage), and a gate terminal of the range mode transistor 1107 is coupled to receive the mode select signal 1102. Thus, when the mode select signal 1102 is high, the range mode enable transistor 1102 is switched on, coupling the source terminals of transistors 604 and 614 to the ground reference voltage (plus a V_t drop through transistor 1107) and enabling the upper and lower bound compare circuits within the range compare circuit 1105. When the mode select

signal 1102 is low, the range mode enable transistor 1107 is switched off (i.e., little or no current flows between the drain and source of transistor 1107), disabling the upper and lower bound compare circuits from asserting the GT_i/LT_i signals. As indicated by dashed lines 1104, 1106, 1114 and 1116, the GT_i and LT_i outputs of the range compare circuit 1105 are alternatively
5 coupled to a match line 127 or output to a more significant CAM cell in a chain of ranging or multi-mode CAM cells. In an alternative embodiment, switches (e.g., transistor switches or logic gates responsive to a select signal) or fuses may be used to couple the GT_i and LT_i outputs to (or decouple the GT_i and LT_i outputs from) the match line 127 in a run-time or production-time configuration operation.

The ternary compare circuit 1109 includes two pairs of compare transistors 1121/1123 and 1125/1127, mask transistor 1129, and ternary mode enable transistor 1108. Transistors 1121, 1123, 1129, and 1108 are coupled in series between a ground reference voltage and match line 127. Gate terminals of the compare transistors 1121 and 1123 are coupled to receive, respectively, a comparand value C from a comparand line and a complemented data value \bar{D} from the storage element 1150. A gate terminal of the mask transistor 1129 is coupled to receive a complemented mask value \bar{M} from storage element 1160, and a gate terminal of the ternary mode enable transistor 1108 is coupled to receive, via inverter 1110, a complemented version of the mode select signal 1102. Compare transistors 1125 and 1127 are coupled in series between the match line 127 and mask transistor 1129, such that transistors 1125, 1127, 1129 and 1108
20 also form a series path between the match line and a ground reference voltage. The gate terminals of the compare transistors 1125 are coupled to receive, respectively, a complemented comparand bit \bar{C} from a comparand line, and a data value D from the storage element 1150.

When the mode select signal 1102 is high, the output of inverter 1110 will be low, switching off transistor 1108. In this state, the ternary compare circuit is disabled from affecting

the state of the match line 127, regardless of the values stored in the storage elements 1150 and 1160. When the mode select signal 1102 is low, the output of inverter 1110 will be high, switching on transistor 1108 and enabling the ternary compare circuit.

When the ternary compare circuit is enabled and a logic '1' is stored in the storage
5 element 1160, the complemented match signal \overline{M} will be low, switching off transistor 1129 and preventing the ternary compare circuit from affecting the state of the match line 1127, regardless of the relative states of data and comparand values. That is, the comparison operation is masked such that the ternary CAM cell effectively stores a "don't care state." When the ternary compare circuit is enabled and a logic '0' is stored in the storage element 1160, the compare transistor pairs 1121/1123 and 1125/1127 will complete a path between the match line and the ground reference potential (thereby pulling the match line low to signal a mismatch condition) if the data value stored in the storage element 1150 does not match an incoming comparand value. For example, if the stored data value is a logic '1' and the comparand value is a logic '0' (a first possible mismatch condition) the complemented comparand signal \overline{C} and the data signal D will both be high, switching on transistors 1125 and 1127. If the stored data value is a logic '0' and the comparand value is a logic '1' (a second possible mismatch condition), the comparand signal C and the complemented data signal will both be high, switching on transistors 1121 and 1123.

Note that the masking function of the ternary compare circuit is achieved by using mask transistor 1129 to interrupt the path between the pairs of compare transistors and the ground
20 reference. In alternative embodiments, the mask transistor may instead be disposed between the ground reference and the ternary mode enable transistor 1108, or between the match line 1127 and the pairs of compare transistors 1121/1123 and 1125/1127. Also, masking techniques other than path interruption may be used in alternative embodiments, including without limitation, disabling the data value from being received by transistors 1123 and 1127 (e.g., switching off

transistors 1123 and 1127 if the mask value is a logic '1'), and/or disabling the comparand value from being received by transistors 1121 and 1125 (e.g., switching off transistors 1121 and 1125 if the mask value is a logic '1'). Such masking techniques are described in WIPO International Publication Number WO 99/23663. In other embodiments, circuitry to implement the masking function may be omitted altogether so that the mode select signal 1102 effectively switches the multi-mode CAM cell 1103 between binary mode (rather than ternary mode) and ranging mode. Also, multiple compare circuits may be included within each multi-mode CAM cell 1103 (e.g., to enable pipelined or parallel compare operations), and the match signal output from the compare circuit 1109 may be combined in a NAND structure, as described above, instead of the wired NOR structure shown in Fig. 11.

Fig. 12 illustrates a mode configuration circuit 1205 that may be used to set the operating modes of multi-mode CAM cells within an exemplary ranging CAM array 1201. The ranging CAM array 1201 includes Y rows of CAM cells 1207(0)-1207(Y-1), each including a number of distinct sets of ternary CAM cells (T) and a number of distinct sets of multi-mode CAM cells (MM). Each set of ternary and multi-mode CAM cells may include any number of CAM cells and may be disposed at any location within a CAM row. Also, while each CAM row 1207 is shown as having ternary and multi-mode CAM cells disposed in the same order and location as in the other rows, the relative order and locations of the sets of ternary and multi-mode CAM cells may vary from row to row in alternative embodiments. Sets of dedicated ranging CAM cells may be included in the CAM array 1201 instead of or in addition to the sets of ternary CAM cells in alternative embodiments, and, in yet other embodiments, all the CAM cells in the CAM array 1201 may be multi-mode CAM cells.

In the embodiment of Fig. 12, the mode configuration circuit 1205 outputs a respective set of mode select signals 1102(0)-1102(Y-1) to each of the rows of CAM cells 1207(0)-1207(Y-

1). Each mode select signal 1102 in a given set of signals is coupled to the mode select input of each multi-mode CAM cell in a given set of multi-mode CAM cells as shown by inset 1210.

Accordingly, the state of each individual mode select signal selects the operating mode for a respective set of multi-mode CAM cells within the CAM array 1201, thereby enabling each set of multi-mode CAM cells to be operated in either ranging mode or ternary mode (or binary mode) regardless of the operating mode of the other sets of multi-mode CAM cells. In one embodiment, the mode configuration circuit 1205 includes storage elements (e.g., registers) that may be programmed by a host processor to set the state of the individual mode select signals 1102, and therefore the operating modes of the various multi-mode CAM cells within the CAM array 1201. In an alternative embodiment, the mode configuration register includes one-time programmable fuses that may be blown, for example at production time, to permanently set the state of the individual mode select signals and establish a particular configuration for the CAM array 1201. In alternative embodiments, a single set of mode select signals may be provided (i.e., instead of a separate set of signals per row of CAM cells) and used to set the operating modes for the CAM array such that each row of CAM cells 1207(0)-1207(Y-1) has the same configuration. Also, instead of multiple mode select signals, a single mode select signal may be input to each multi-mode CAM cell in the CAM array 1201 to set the operating mode for each multi-mode CAM cell in the CAM array 1201. Further, the mode select signal may be provided from an external source (e.g., a host processor or other device, or by strapping) instead of from the mode configuration circuit 1205.

Fig. 13 illustrates an embodiment of a ranging CAM array 1301 that includes multiple independently addressable and searchable blocks of CAM cells 1305(0)-1305(K-1). Such storage blocks are referred to herein as CAM blocks. Each row of CAM cells within a given CAM block 1305 includes multiple row segments, S_1 - S_z , that may be alternately configured to

store an individual CAM word or a portion of a CAM word that spans two or more row segments. The storage width and depth of the CAM block is effectively determined by the configuration of the row segments. Assume, for example, that a given CAM block 1305 includes Y rows of CAM cells, Z row segments per row, and W CAM cells per row segment. If the row segments within the CAM block 1305 are configured to store CAM words that span all Z row segments, then the storage width of the CAM block becomes $W \times Z$ and the storage depth of the CAM block becomes Y. Alternatively, if the row segments within the CAM block 1305 are configured to store a single CAM word per row segment, then the storage width of the CAM block becomes W and the storage depth of the CAM block becomes $Z \times Y$. Other width/depth configurations between the $W \times Z$ by Y and W by $Z \times Y$ configurations may also be achieved (e.g., $W \times (Z/2)$ by $Y \times (Z/2)$).

To add further flexibility in the configuration of the CAM blocks 1305, each of the individual row segments within each CAM block may include different types of CAM cells including one or more sets of ternary (or binary) CAM cells, ranging CAM cells and/or multi-mode CAM cells. Block configuration circuits 1307(0)-1307(K-1) are provided to control the respective width/depth configurations of the CAM blocks 1305 and to set the operating modes of multi-mode CAM cells within the CAM blocks 1305. In the embodiment of Fig. 13, for example, each block configuration circuit 1307 outputs one or more sets of mode select signals (MSEL) to a corresponding one of the CAM blocks 1305 in the manner described in reference to Fig. 12, and each block configuration circuit 1307 also outputs a width/depth configuration signal (CFG) to set the width/depth configuration of the corresponding CAM block 1305. As in the CAM array 1201 of Fig. 12, the individual sets of multi-mode CAM cells within a given CAM block 1305 may each receive a dedicated mode select signal, or mode select signals may be shared between some or all of the sets of multi-mode CAM cells.

Fig. 14 illustrates a multi-mode CAM 1401 cell that may be switched between an implicit range mode and a binary mode. The CAM cell 1401 includes a storage element 1405, a range compare circuit 1407 and a binary compare circuit 1409. The storage element 1405 is implemented by a pair of back-to-back coupled inverters and is used, alternatively, to store a boundary value B for use in the range compare circuit 1407 or a data value D for use by the binary compare circuit 1409. As discussed in reference to Fig. 6, numerous other types of storage cells may be used to implement storage element 1405 in alternative embodiments.

The range compare circuit 1407 includes transistors 602, 604, 606 and 608 each coupled as described in reference to Fig. 6 except that the source terminal of transistor 604 is coupled to the drain terminal of a range mode enable transistor 1107, and the gate terminals of transistors 602 and 608 are coupled to the output of a multiplexer 1411 instead of directly to a comparand line. The range mode enable transistor 1107 is coupled between the source terminal of transistor 604 and a reference voltage as described in reference to Fig. 11, and is responsive to a mode select signal 1102 to either enable or disable the range compare circuit 1407. More specifically, when the mode select signal 1102 is high, transistor 1107 is switched on, enabling the range compare circuit 1407. When the mode select signal is low, transistor 1107 is switched off, disabling the range compare circuit 1407.

The multiplexer 1411 is responsive to a level select signal 1404 to select between a comparand value C and its complement \bar{C} . When the level select signal 1404 is in an upper bound state (e.g., high), the comparand value C is selected to be output to the gate terminals of transistors 602 and 608, and the boundary value B is treated as the complement of an upper bound value. That is, the range compare circuit 1407 compares the comparand value C and the upper bound value as described above in reference to Fig. 6, asserting a GT_i signal at the drain terminal of transistor 602 if the comparand C is greater than the upper bound value, or if the

comparand C is equal to the upper bound value and a GT_{i-1} signal is asserted at the source terminals of transistors 606 and 608.

When the level select signal 1404 is in a lower bound state (e.g., low), the complemented comparand value \bar{C} is selected to be output to the gate terminals of transistors 602 and 608, and the boundary value B is treated as a lower bound value. The range compare circuit 1407 compares the complemented comparand value \bar{C} and the lower bound value as described above in reference to Fig. 6, asserting an LT_i signal at the drain of the transistor 602 if the comparand is less than the lower bound value, or if the comparand is equal to the lower bound value and a LT_{i-1} signal is asserted at the source terminals of transistors 606 and 608. Thus, the GT/LTi signal output from the range compare circuit 1407 represents either a GT signal or a LT signal according to the state of the level select signal 1404. Similarly, the GT/LTi-1 signal input from a less significant multi-mode CAM cell (or dedicated ranging CAM cell) is deemed to be either a GT signal or a LT signal according to the state of the level select signal 1404. By this arrangement, the level select signal may be used to switch the implicit range compare function of the multi-mode CAM cell 1403 to either a greater-than compare or a less-than compare. As with the GTi and LTi outputs of the CAM cell described in reference to Fig. 11, the GT/LTi output of the multi-mode CAM cell 1403 may either be coupled to the match line 127 as shown by dashed line 1414 (e.g., if the multi-mode CAM cell 1403 is the most significant CAM cell in a chain) or coupled to the GT/LT input of a more significant CAM cell in a CAM cell chain as shown by dashed line 1416. Respective level select signals 1404 may be output to individual chains of multi-mode CAM cells 1403 by a configuration circuit similar to the mode configuration circuit of Fig. 1205, or a single level select signal 1404 may be shared by multiple (or all) chains of multi-mode CAM cells 1403 within a CAM array. Also, the block configuration circuits 1307(0)-1307(K-1) of Fig. 13 may additionally output one or more level select signals 1404 to

control the compare function within chains of multi-mode CAM cells 1403 within the configurable blocks 1305(0)-1305(K-1). Alternatively, the level select function (i.e., the function implemented by multiplexer 1411 and level select signal 1404) may be omitted altogether from the multi-mode CAM cell 1403. In that case, one or the other of the comparand values (C or \bar{C}) may be input directly to the gate terminals of transistors 602 and 608 to implement either an upper bound or lower bound compare function. For example, the signal line used to carry the uncomplemented comparand value C may be coupled directly to the gate terminals of transistors 602 and 608, and a complemented upper bound value \bar{U} applied to the gate terminals of transistors 604 and 606, thus implementing an upper bound compare function.. A lower bound compare function may be implemented without circuit change within CAM cell 1403 by inverting the interpretation of the value stored in the storage element 1405 (i.e., interpret the stored value as \bar{L} instead of U, such that a stored '1' = \bar{L} , and a stored '0' = L) and by inverting the interpretation of the comparand signals issued on the signal line used to carry the uncomplemented comparand value (i.e., interpret the comparand value as \bar{C} instead of C (i.e., signal level '0' = C, and signal level '1' = \bar{C})). Such inversions may be accomplished at a system level by a host processor or other device.

Still referring to Fig. 14, the binary compare circuit 1409 includes transistors 1121, 1123, 1125 and 1127 coupled as described in reference to Fig. 11, except that the source terminals of transistors 1123 and 1127 are coupled to the drain terminal of a binary mode enable transistor 1408. A source terminal of the binary mode enable transistor is coupled to a reference voltage (e.g., directly or through a timing control circuit) and a gate terminal of the binary mode enable transistor 1408 is coupled, through an inverter 1410, to the mode select signal 1102. Thus, when the mode select signal is low, transistor 1107 is switched off to disable the range compare circuit

1407 and transistor 1408 is switched on to enable the binary compare circuit 1409. When enabled, the binary compare circuit 1409 pulls the match line 127 to a low logic level if the comparand value does not match the data value stored in storage element 1405 (i.e., if C and \bar{D} are high, or if \bar{C} and D are high). When the mode select signal 1102 is low, transistor 1107 is
5 switched on to enable the range compare circuit 1407 and transistor 1408 is switched off to disable the binary compare circuit 1409.

Fig. 15 illustrates an alternative compare circuit 1500 that may be used in place of the binary compare circuit 1409 of Fig. 14. Compare circuit 1500 includes a first transistor 1505 coupled between a reference potential (e.g., ground) and a match line 127. A gate terminal of the transistor 1505 is coupled to a drain terminal of a second transistor 1501 and to a drain terminal of a third transistor 1503. Source terminals of transistors 1501 and 1503 are coupled to receive a comparand signal and its complement, respectively, while gate terminals of transistors 1501 and 1503 are coupled to receive a complemented data value (e.g., a complement of the data value stored in storage element 1405 of Fig. 14), and an uncomplemented data value, respectively. By this arrangement, if the comparand signal is high and the data value is low (a mismatch condition), transistor 1501 will be switched on, passing the comparand signal (i.e., high logic level signal) to the gate of transistor 1505. Transistor 1505 will be switched on in response to the high logic level signal at its gate terminal, coupling the match line to the reference voltage to indicate the mismatch condition. Similarly, if the comparand is low and the data value is high
20 (also a mismatch condition), transistor 1503 will be switched on, passing the complemented comparand signal (i.e., high logic level signal) to the gate terminal of transistor 1505. Again, transistor 1505 will be switched on, coupling the match line 127 to the reference voltage to indicate the mismatch condition.

Note that compare circuit 1505 may be modified to enable a ternary compare operation, for example, by including an additional transistor in series with transistor 1505 (i.e., between the reference potential and the match line 127) and coupling a gate terminal of the additional transistor to a match value. Alternatively, circuitry responsive to a match value may be used to prevent logic high data and/or comparand signals from being applied to the terminals of transistors 1501 and 1503. In general, any circuit for detecting a mismatch between data and comparand values may be used to implement compare circuits within the binary and ternary CAM cells described herein, and any circuit for preventing a compare circuit from indicating a mismatch condition may be used as a masking circuit within the ternary CAM cells described herein. Moreover, as discussed above, multiple compare circuits may be included within each CAM cell disclosed herein and the match outputs of a set of CAM cells may be coupled in a NAND structure (e.g., pass gates coupled in series between a match line output and a reference voltage) instead of the NOR structure shown, for example, in Figs. 11 and 14. Fig. 16 illustrates a multi-mode CAM cell 1603 according to another alternative embodiment. The multi-mode CAM cell 1603 includes a storage element 1605 and a mode-selectable compare circuit 1607. The storage element 1605 is implemented by a pair of back-to-back coupled inverters and is used, alternatively, to store a boundary value (an upper bound value, U, in this example) for use in a range comparison or a data value D for use in a binary comparison. As discussed in reference to Fig. 6, numerous other types of storage cells may be used to implement storage element 1605 in alternative embodiments.

The mode-selectable compare circuit 1607 includes transistors 1121, 1123, 1125 and 1127 coupled to one another in the manner discussed above in reference to Figs. 11 and 14. The gate terminals of transistors 1123 and 1127 are coupled to receive complementary signals (e.g., D and \bar{D} , respectively) from the storage element 1605, and the gate terminal of transistor 1125

is coupled to receive a comparand value C from a comparand signal line. Transistors 606 and 608 are coupled to one another in the same manner as described above in reference to Fig. 6. The drain terminals of transistors 606 and 608 are coupled to the drain terminal of transistor 1125 and the gate terminals of transistors 606 and 608 are coupled to the gate terminals of transistors 1127 and 1125 respectively. Together, transistors 1125, 1127, 606 and 608 form a range compare circuit in the same manner as transistors 602, 604, 606 and 608, respectively, of Fig. 6.

Sub A1
The mode selectable circuit further includes transistors 1609, 1611 and 1613, each having a gate terminal coupled to receive the mode select signal 1102. A drain terminal of transistor 1609 is coupled to the source terminals of transistors 606 and 608, and a source terminal of transistor 1609 is coupled to receive an input GT signal (e.g., GT_{i-1}). A drain terminal of transistor 1611 is coupled to a gate terminal of transistor 1121, and a source terminal of transistor 1611 is coupled to receive a complemented comparand signal from a comparand signal line. A drain terminal of transistor 1613 is coupled to the match line 127, and a source terminal of transistor 1613 is coupled to the drain terminals of transistors 1121 and 1125. In the embodiment of Fig. 16, transistors 1613 and 1515 are PMOS (Positively doped, Metal Oxide Semiconductor) devices that are switched on (i.e., enter a conducting state) when the mode select signal 1102 is low, and switched off when the mode select signal 1102 is high. Transistor 1609 is an NMOS (Negatively doped, Metal Oxide Semiconductor) device that is switched on when the mode select signal 1102 is high, and switched off when the mode select signal is low.

When the mode select signal 1102 is high, a range mode of operation is selected for the CAM cell 1603, with transistor 1609 being switched on to allow a GT input signal (i.e., GT_{i-1}) to be received at the source terminals of transistors 606 and 608, transistor 1611 being switched off to prevent the complemented comparand signal from being received at the gate terminal of

transistor 1121, and transistor 1613 being switched off to prevent the GT output (i.e., GT_i) from affecting the state of the match line 127. If the CAM cell 1603 is the most significant CAM cell in a chain of range CAM cells (including a chain of multi-mode CAM cells in range mode), then transistor 1613 may be omitted and the drain of transistor 1125 coupled directly to the match line 127 as shown by dashed line 1606. If the CAM cell 1603 is not the most significant CAM cell in a chain of range CAM cells, the GT_i signal is output to a next more significant range CAM cell (or multi-mode CAM cell in range mode) as shown by dashed line 1608. Although the range compare function shown in Fig. 16 corresponds to an upper bound compare function (i.e., determining whether a comparand value exceeds an upper bound value), a lower bound compare function may be implemented without circuit change by applying a complemented version of the comparand value to the comparand lines (such that \bar{C} is received at the gate terminals of transistors 608 and 1125 instead of C) and by storing a complemented version of the lower bound data within storage element 1605 (such that L is received at the gate of transistor 1127 instead of \bar{U}). Alternatively, a multiplexer responsive to a level select signal may be provided to select between complemented and uncomplemented versions of the comparand value (i.e., between C and \bar{C}) as in the embodiment of Fig. 14.

Still referring to Fig. 16, when the mode select signal 1102 is low, a binary mode of operation is selected, with transistor 1609 being switched off to prevent the incoming GT signal (GT_{i-1}) from being received at the source terminals of transistors 606 and 608, transistor 1611 being switched on to conduct the complemented comparand value \bar{C} to the gate terminal of transistor 1121, and transistor 1613 being switched on to allow the output of the binary compare circuit formed by transistors 1121, 1123, 1125 and 1127 to pull the match line 127 low if the data value stored in the storage element 1605 does not match the incoming comparand value. Note that, in an alternative embodiment, transistor 1611 may be disposed in the path between the

storage element 1605 and the gate of transistor 1123 instead of in the path between the comparand signal line and the gate of transistor 1121. Further, pass gates responsive to the mode select signal 1102 (i.e., such as transistor 1611) may be disposed in the paths to the gates of both transistors 1121 and 1123. Also, in an alternative embodiment, transistors 1611 and 1613 may
5 be NMOS devices instead of the PMOS devices, with signal inverters coupled between the mode select input of the CAM cell 1603 and the gate terminals of the transistors 1613 and 1611.

System Structure and Operation

Fig. 17 illustrates a system 1700 that includes a host processor 1703 (e.g., general purpose processor, digital signal processor, network processor, application-specific integrated circuit (ASIC), etc.) and a ranging CAM device 1701 according to one of the embodiments described herein. The system device may be, for example, a network switch or router, or any other type of device in which the compare capability of the CAM device 1701 may be useful.

The host processor 1703 issues addresses, comparands, and instructions to the ranging CAM device 1701 via the address, comparand and instruction buses, respectively (i.e., ABUS 113, CBUS 115 and IBUS 111), and receives status and other information from the ranging CAM device 1701 via a result bus 149 (RBUS). In the embodiment of Fig. 17, the flag signal (e.g., match flag, multiple match flag, full flag, etc.) is output directly to the host processor 1703, however, the flag signal (or flag signals) may alternatively or additionally be output to the host processor 1703 via the result bus 149. The CAM index (i.e., the index generated by the priority
20 encoder 105 of Fig 1) may be output to an associated storage and/or to the host processor 1703. Also, in alternative embodiments, one or more of the buses (e.g., ABUS, CBUS, IBUS, ADS) may be omitted and the corresponding information time multiplexed onto another of the buses. Further, the ranging CAM device 1701 and host processor 1703 may be implemented in distinct integrated circuits (ICs) and packaged in distinct IC packages, or in a single IC (e.g., in an ASIC,

system-on-chip, etc.), or in an IC package that includes multiple ICs (e.g., a multi-chip package, paper thin package, etc.).

Fig. 18 illustrates the operation of the host processor 1703 of Fig. 17 according to one embodiment. At block 1801 the host processor receives an assignment rule that indicates an assignment of destination or classification information (e.g., destination address, destination port, class ID, etc.) to incoming packets having a particular value or range of values within one or more fields (e.g., fields within the packet prefix). At decision block 1803, the host processor determines whether the assignment rule specifies assignment of the destination or classification information to a range of packet values. If so, then a range definition may be used to express the assignment rule, and at decision block 1805 the processor determines whether a range field (i.e., chain of ranging CAM cells or multi-mode CAM cells) having a sufficient data storage width and CAM word offset is available within the ranging CAM device to store the range definition. If a chain of multi-mode CAM cells is available to store the range definition, then at block 1807, the host processor issues an instruction to the ranging CAM device identifying the chain of multi-mode cells and instructing the ranging CAM device to set the multi-mode CAM cells to operate in range mode. If the chain of multi-mode CAM cells is already in range mode (e.g., by default at system startup), or if the CAM cells are dedicated range CAM cells rather than multi-mode CAM cells, then block 1807 may be omitted. At block 1809, the host processor issues an write instruction to the CAM device to store an upper bound value in a given entry (entry_i) within the ranging CAM array of the ranging CAM device. The row address of the entry may be included in the write instruction, or the write instruction may reference an address value stored within the ranging CAM device, such as a register that stores an address of an available row of CAM cells within the ranging CAM device. At block 1811, the host processor issues a write instruction to the CAM device to store a lower bound value in entry_i within the CAM device,

thus effecting storage of an explicit range definition within the CAM device. Note that the write operations to store the upper and lower bound values are shown in parallel in blocks 1809 and 1811 to indicate that a single write command (or request or instruction) may be used to effect storage of the upper and lower bound values. Also, storage of either the upper bound value or the lower bound value may be omitted in the case of an implicit range definition.

At block 1813, the host processor issues a write instruction to the ranging CAM device to store data bits within binary or ternary CAM cells included within entry_i of the CAM array. At block 1815, the host processor issues a write instruction to the ranging CAM device to store mask bits within ternary CAM cells included within entry_i of the CAM array. Note that blocks 1813 and 1815 may be omitted in the event that the range definition (or definitions) constitutes the entire CAM word stored within entry_i. Also, if there are no mask bits to be stored in ternary CAM cells within entry_i, or there are no ternary CAM cells within the CAM array at entry_i (e.g., binary CAM cells instead of ternary CAM cells), block 1815 may be omitted. In any event, a single write operation, or multiple write operations, may be used to write the upper bound, lower bound, data and/or mask values into the CAM array at entry_i. Returning to decision blocks 1803 and 1805, if no range assignment is needed or desirable to express the assignment rule received in block 1801, or if there is no available range field within the ranging CAM device, then at blocks 1817 and 1819, data bits and, if necessary, mask bits are stored within CAM array entry_i (or multiple entries of the CAM array) to implement the assignment rule.

Although the invention has been described with reference to specific exemplary embodiments thereof, it will be evident that various modifications and changes may be made thereto without departing from the broader spirit and scope of the invention as set forth in the appended claims. The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense.